Basin Area to Stream Points vs. Distance from Divide to Stream Point

Marianna Alexandersdottir
NWIFC
March 17, 2002

This analysis is intended to evaluate the relationship of basin areas to the stream distance from the divide to the point on the stream characterizing the basin area, i.e. the point of discontinuous water (P_d) and of continuous flowing water (P_p).

Distribution of distances.

The distribution of the distances, like that of the basin areas, is skewed to the right, as shown in Figure 1 below. The data are more normally distributed when transformed using the natural logarithm as also shown in Figure 1 below. The QQ plots show that the log transformed data depart slightly at the tails from a natural distribution, but not severely. For the purposes of this analysis the log-transformed data are used. The mean and median distances, with standard errors for the mean are shown in Table 1 for the data on the original scale and for the log-transformed data. The 90% confidence interval for each cooperator is shown in Figure 2.

Relationship between distance to stream points P_d and P_p and basin areas

Linear regressions on the log-transformed data were used to evaluate the relationship between the basin areas measured to the point in the stream where discontinuous water started (P_d) and to where continuous flowing or perennial water was detected (P_p). The data were collected by cooperating agencies and are grouped within two major regions, the Westside where the default basin area is 52 acres and the Eastside where the default basin area is 300 acres. Regressions were done for each cooperator separately (Table 2).

In order to determine whether there were any differences among the datasets collected by the cooperating agencies Analysis of Covariance was used to test the hypotheses that there was no difference in the slope and intercepts of the regressions (Table 2). The model is as follows:

$$\ln(Px) = \ln(Ax) + Cooperator + \ln(Ax) \bullet Cooperator$$

where $\ln(P_x)$ is the log-transformed distance to a stream point and $\ln(A_x)$ is the log-transformed basin area defined by that point. If there is a significant interaction term this indicates there are differences in the slopes, while a significant main effect for cooperator indicates differences in the intercepts for the separate regressions.

For the basin area measured to P_d neither term was significant (Cooperator or interaction) in either region (Table 3), while these were both significant in the 300 acre default area region for the area measured to point P_p. This result indicates that the regressions are coincident for the P_d distance in both regions, but only in the regions with 52 acre defaults for the P_p distance (Table 3). The sample sizes are smallest for the cooperators in the eastside (300 acre) region and some points are very influential in the regressions (Figure 3), which may be resulting in this significant

difference among cooperators. When cooperators are combined and regressions compared between regions (52 vs 300 acre default), no significant difference was found (Table 3, Figure 4).

The relationship on the original scale between distance to a stream point and basin area is a power function,

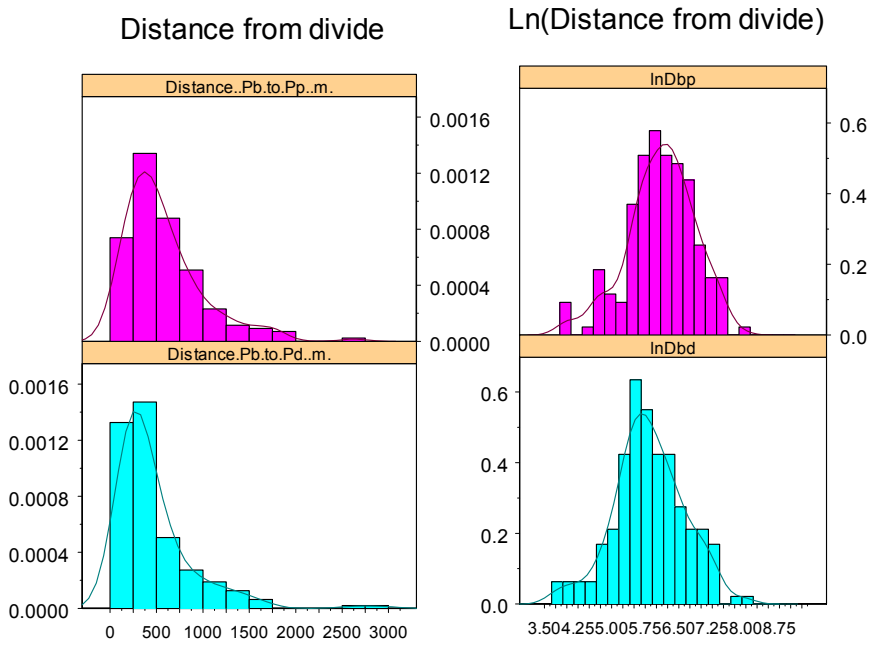
$$\textit{Distance to } P_x = \alpha A_x^b$$

where α is the exponential of the intercept of the regression and b is the slope of the regression (Figure 5). This function shows the decreasing rate at which the distance to the point of discontinuous water lengthens with increasing basin size.

Regressions of the log-transformed distance to P_d as a function of basin area A_d for each region separately and combining all data, were significant with R^2s ranging from 66-85%, and an R^2 of 70% for the statewide data (Table 4). Figure 6 shows the resulting regression with 90% prediction intervals.

Similar relationships exist for distance to perennial water P_p with basin area in the 52 and 300 acre regions and for all data combined (Table 4). But given the small sample sizes and differences between cooperators in the 300 acre default region, this relationship should be used with caution. In order to better define the relationship for P_p additional data collection would be necessary.

A. Distributions of distances and distances transformed using natural logarithm.



B. QQ plots for data on original scale and log-transformed data.

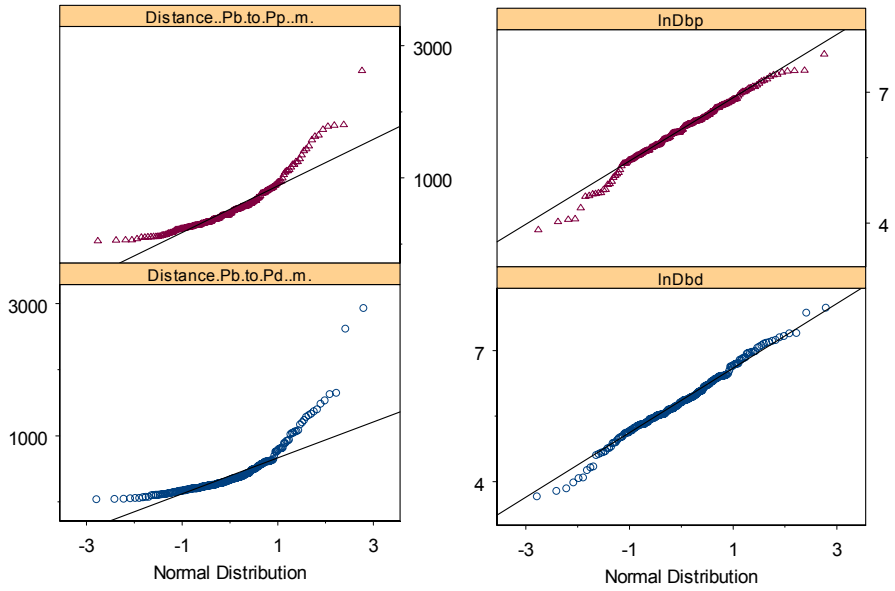
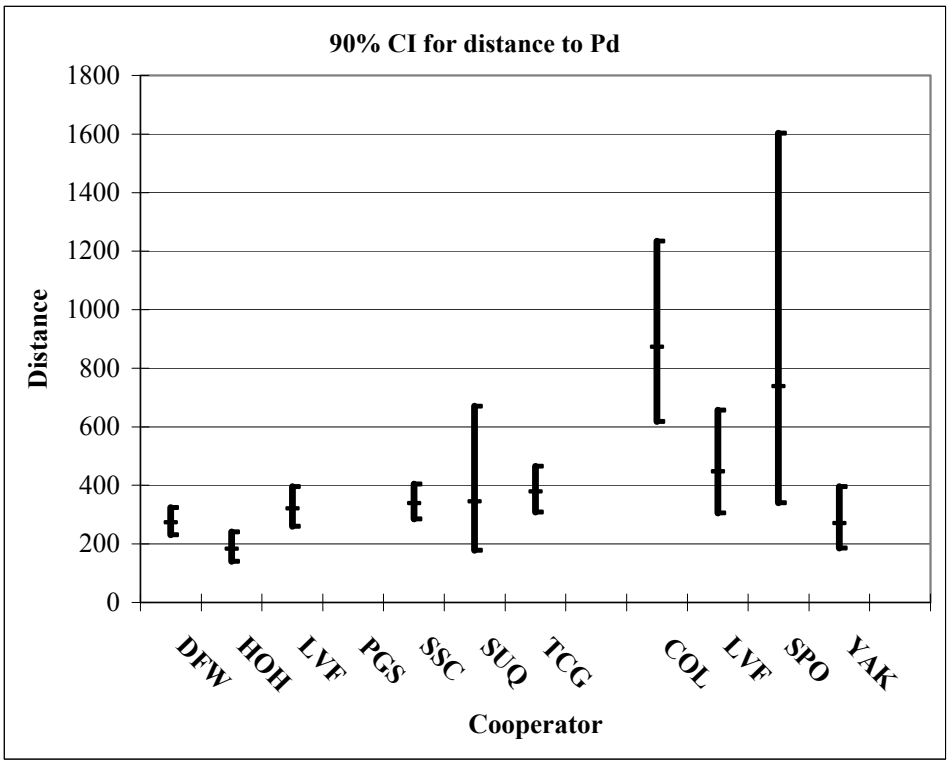


Figure 1. Distribution (A) and QQ plots (B) of measurements of distance from divide to points (P_d and P_p) defining basin areas.

Table 1. Statistics for distances to Pd and Pp by cooperator and region (default basin areas).

Region	Cooperator	Original Scale				Log-Transformed			
		N	Mean	Median	SE	N	Mean	Median	SE
Distance to Pd									
52	DFW	22	306.4	272.0	36.4	22	5.6132	5.6038	0.0978
	HOH	21	229.1	194.0	45.9	20	5.2134	5.3530	0.1559
	LVF	22	376.6	299.0	54.9	22	5.7707	5.7008	0.1222
	PGS	1	210.0	210.0		1	5.3471	5.3471	
	SSC	23	383.0	309.0	43.0	23	5.8282	5.7333	0.1013
	SUQ	5	417.2	258.0	124.5	5	5.8442	5.5530	0.3114
	TCG	55	530.6	395.0	54.1	55	5.9379	5.9793	0.1224
52		149	403.5	293.0	25.8	148	5.7427	5.7020	0.0609
300	COL	10	1,034.4	830.0	215.9	10	6.7728	6.6864	0.1886
	LVF	13	633.3	360.0	203.3	13	6.1049	5.8861	0.2142
	SPO	6	949.8	983.5	218.5	6	6.6049	6.8893	0.3843
	YAK	12	323.4	350.5	44.8	12	5.6030	5.8563	0.2102
300		41	686.8	441.0	97.5	41	6.1941	6.0890	0.1336
Distance to Pp									
52	DFW	17	650.5	561.0	87.0	17	6.3642	6.3288	0.1160
	HOH	18	311.8	244.0	56.2	18	5.5207	5.4980	0.1608
	LVF	22	433.6	351.5	63.9	22	5.8934	5.8594	0.1283
	PGS	1	214.0	214.0		1	5.3658	5.3658	
	SSC	23	649.5	577.0	78.4	23	6.3354	6.3578	0.1119
	SUQ	5	495.2	430.0	123.2	5	6.0733	6.0648	0.2602
	TCG	49	550.6	395.0	64.0	49	5.9415	5.9793	0.1347
52		135	524.6	422.0	33.0	135	5.9985	6.0450	0.0662
300	COL	9	1,110.3	1,106.0	229.9	9	6.8578	7.0085	0.1939
	LVF	12	571.6	507.5	70.7	12	6.2699	6.2297	0.1187
	SPO	5	1,114.4	1,163.0	243.2	5	6.8751	7.0588	0.2949
	YAK	12	562.7	513.0	86.0	12	6.1821	6.2404	0.1798
300		38	767.8	592.5	81.1	38	6.4610	6.3837	0.1006



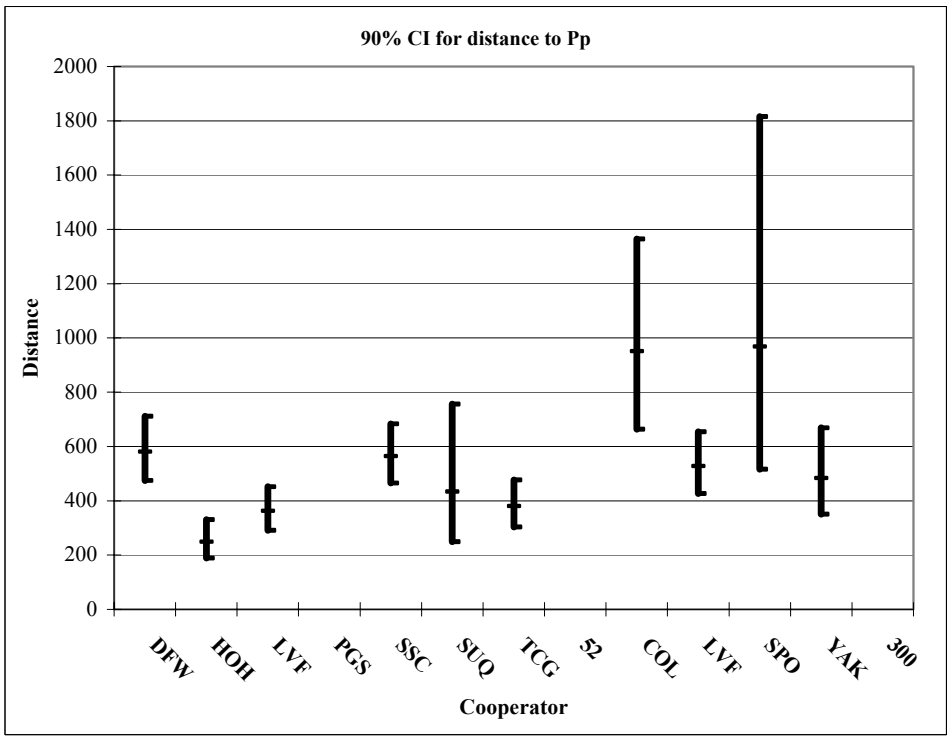


Figure 2. 90% confidence intervals for distance to discontinuous (P_d) and perennial water (P_p) by cooperator.

Table 2. Results of regressions of log-transformed distance to stream points P_d and P_p on log-transformed basin areas to those points.

Default Basin Area	Cooperator		ln(Distance to Pd)=f(Ad)				ln(Distance to Pp)=f(Ap)			
			Intercept	Slope	df	R^2	Intercept	Slope	df	R^2
52	DFW	Estimate	4.82	0.44	20	52.7%	5.22	0.36	15	72.6%
		SE	0.17	0.09			0.18	0.06		
		Lower 95% CI	4.46	0.25			4.83	0.25		
		Upper 95% CI	5.18	0.63			5.61	0.48		
		t-statistic	27.73	4.94			28.43	6.58		
		p-value	0.0000	0.0001			0.0000	0.0000		
	HOH	Estimate	4.90	0.41	16	79.7%	4.99	0.42	15	87.5%
		SE	0.09	0.05			0.08	0.04		
		Lower 95% CI	4.71	0.31			4.83	0.34		
		Upper 95% CI	5.08	0.52			5.15	0.50		
		t-statistic	56.47	8.23			65.64	10.64		
		p-value	0.0000	0.0000			0.0000	0.0000		
	LVF	Estimate	5.01	0.37	20	60.2%	5.11	0.29	20	64.9%
		SE	0.15	0.06			0.15	0.05		
		Lower 95% CI	4.69	0.23			4.80	0.19		
		Upper 95% CI	5.33	0.50			5.41	0.38		
		t-statistic	32.64	5.72			34.93	6.31		
		p-value	0.0000	0.0000			0.0000	0.0000		
	PGS	Estimate	5.35	-			5.37	-		
		SE								
		Lower 95% CI								
		Upper 95% CI								
		t-statistic								
		p-value								
	SSC	Estimate	5.40	0.32	15	27.7%	5.36	0.41	17	53.5%
		SE	0.20	0.12			0.24	0.09		
		Lower 95% CI	4.97	0.06			4.86	0.23		
		Upper 95% CI	5.84	0.57			5.86	0.60		
		t-statistic	26.52	2.67			22.65	4.66		
		p-value	0.0000	0.0174			0.0000	0.0002		
	SUQ	Estimate	4.58	0.53	3	82.4%	4.80	0.47	3	81.2%
		SE	0.31	0.12			0.32	0.11		
		Lower 95% CI	3.59	0.15			3.78	0.12		
		Upper 95% CI	5.58	0.91			5.81	0.83		
		t-statistic	14.70	4.45			15.04	4.27		

Table 2. Results of regressions of log-transformed distance to stream points P_d and P_p on log-transformed basin areas to those points.

Default Basin Area	Cooperator		ln(Distance to Pd)=f(Ad)				ln(Distance to Pp)=f(Ap)			
			Intercept	Slope	df	R^2	Intercept	Slope	df	R^2
		p-value	0.0007	0.0211			0.0006	0.0236		
		Estimate	4.68	0.47	53	62.0%	4.65	0.46	47	56.4%
		SE	0.15	0.05			0.18	0.06		
	TCG	Lower 95% CI	4.38	0.37			4.28	0.34		
		Upper 95% CI	4.99	0.58			5.03	0.57		
		t-statistic	30.67	9.44			25.17	7.94		
		p-value	0.0000	0.0000			0.0000	0.0000		
300		Estimate	4.82	0.45	8	81.7%	4.86	0.44	7	90.6%
		SE	0.31	0.07			0.23	0.05		
	COL	Lower 95% CI	4.10	0.29			4.31	0.33		
		Upper 95% CI	5.55	0.61			5.41	0.56		
		t-statistic	15.32	6.42			20.90	8.86		
		p-value	0.0000	0.0002			0.0000	0.0000		
		Estimate	4.63	0.51	11	89.4%	5.64	0.16	9	13.9%
		SE	0.16	0.05			0.42	0.10		
	LVF	Lower 95% CI	4.27	0.40			4.70	(0.06)		
		Upper 95% CI	4.98	0.62			6.58	0.38		
		t-statistic	28.58	10.12			13.57	1.62		
		p-value	0.0000	0.0000			0.0000	0.1400		
		Estimate	4.88	0.41	4	83.1%	5.31	0.29	2	77.2%
		SE	0.38	0.08			0.49	0.09		
	SPO	Lower 95% CI	3.83	0.18			3.21	(0.08)		
		Upper 95% CI	5.92	0.63			7.41	0.65		
		t-statistic	12.97	5.05			10.88	3.34		
		p-value	0.0002	0.0072			0.0083	0.0790		
		Estimate	4.58	0.42	10	70.3%	4.32	0.49	10	91.9%
		SE	0.23	0.08			0.17	0.04		
	YAK	Lower 95% CI	4.08	0.24			3.93	0.39		
		Upper 95% CI	5.09	0.61			4.71	0.58		
		t-statistic	20.15	5.20			24.88	11.23		
		p-value	0.0000	0.0004			0.0000	0.0000		

Table 3. Results for Analysis of Covariance testing the hypothesis of equal intercepts and slopes among cooperators within default basin areas and among default basin areas for cooperators combined.

Source	df	SS	F	p-value
Distance to Pd				
52 Acre Default Basin Area, compare cooperators				
Ln (Basin Area)	1	16.32	81.08	0.0000
Cooperator	5	1.77	1.76	0.1256
Interaction	5	0.55	0.54	0.7444
Error	127	25.56		
300 Acre Default Basin Area, compare cooperators				
Ln (Basin Area)	1	15.91	154.85	0.0000
Cooperator	3	0.09	0.29	0.8324
Interaction	3	0.13	0.44	0.7279
Error	33	3.39		
Combined data, compare basin areas				
Ln (Basin Area)	1	63.44	337.47	0.0000
Default Basin Area	1	0.55	2.91	0.0899
Interaction	1	0.27	1.44	0.2315
Error	177	33.28		
Distance to Pp				
52 Acre Default Basin Area, compare cooperators				
Ln (Basin Area)	1	16.82	77.96	0.0000
Cooperator	5	1.74	1.61	0.1627
Interaction	5	1.21	1.12	0.3528
Error	117	25.25		
300 Acre Default Basin Area, compare cooperators				
Ln (Basin Area)	1	7.34	93.41	0.0000
Cooperator	3	0.92	3.90	0.0190
Interaction	3	1.08	4.57	0.0100
Error	28	2.20		
Combined data, compare basin areas				
Ln (Basin Area)	1	32.40	140.90	0.0000
Default Basin Area	1	0.01	0.05	0.8202
Interaction	1	0.05	0.21	0.6490
Error	162	37.25		

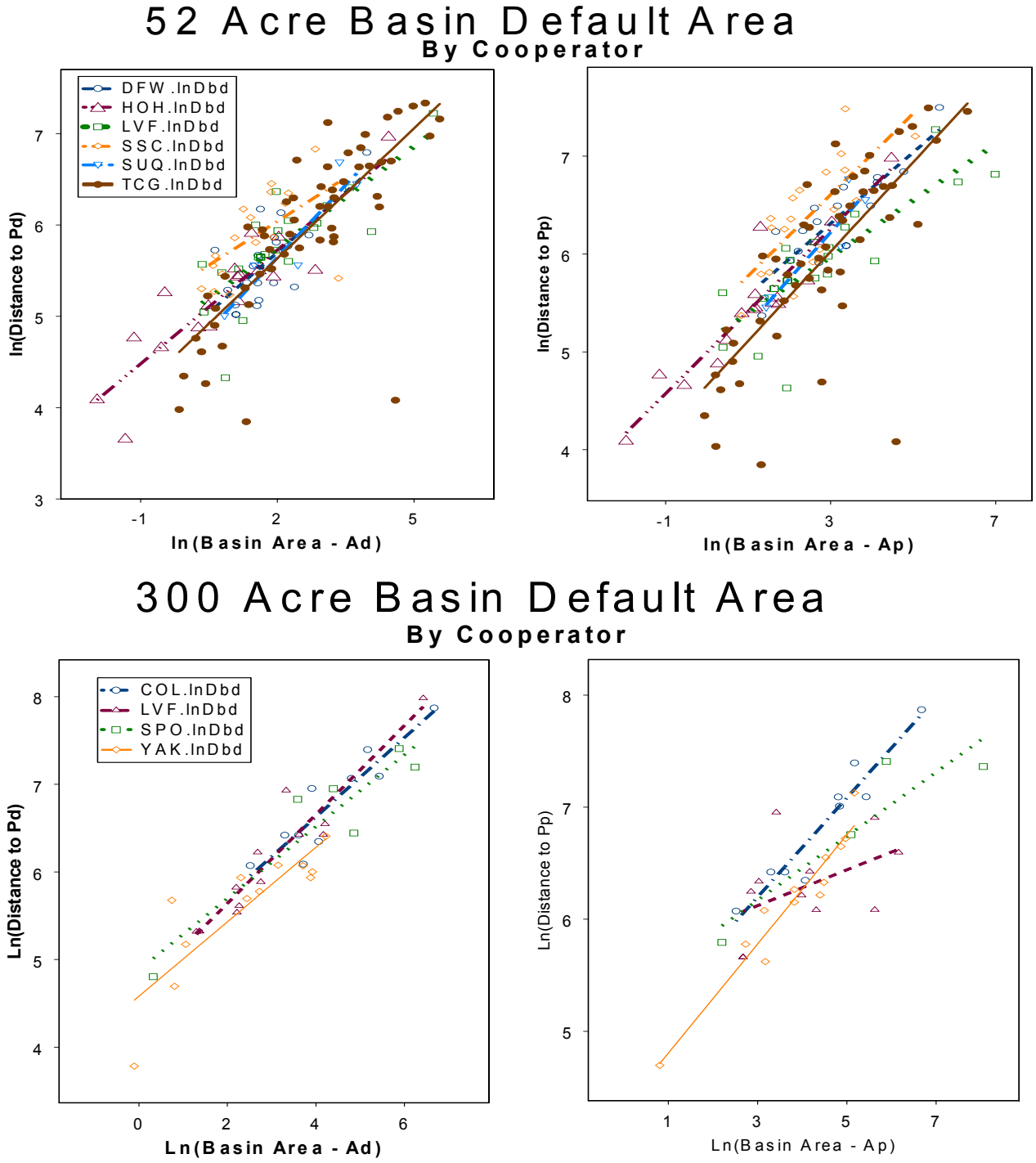


Figure 3. **Scatter plot of basin area and distance from divide to Pd and Pp for regions with default basin areas 52 and 300 default basin area showing scatter for each cooperator separately. Both measurements are log-transformed.**

Table 4. Estimates of intercept and slope of regression for region wide and statewide combined data.

	Estimate	SE	t-statistic	p-value	R2
Distance to P_d					
52 acre					
Intercept	4.889	0.072	67.50	0.00000	63.4%
Slope	0.424	0.029	14.50	0.00000	
300 acre					
Intercept	4.652	0.116	39.82	0.00000	85.5%
Slope	0.461	0.032	14.15	0.00000	
combined					
Intercept	4.863	0.062	78.87	0.00000	69.9%
Slope	0.424	0.022	19.02	0.00000	
Distance to P_p					
52 acre					
Intercept	4.954	0.086	56.97	0.00000	60.2%
Slope	0.391	0.029	13.49	0.00000	
300 acre					
Intercept	4.928	0.192	26.65	0.00000	67.4%
Slope	0.362	0.043	8.38	0.00000	
combined					
Intercept	4.980	0.076	65.61	0.00000	63.6%
Slope	0.370	0.022	16.52	0.00000	

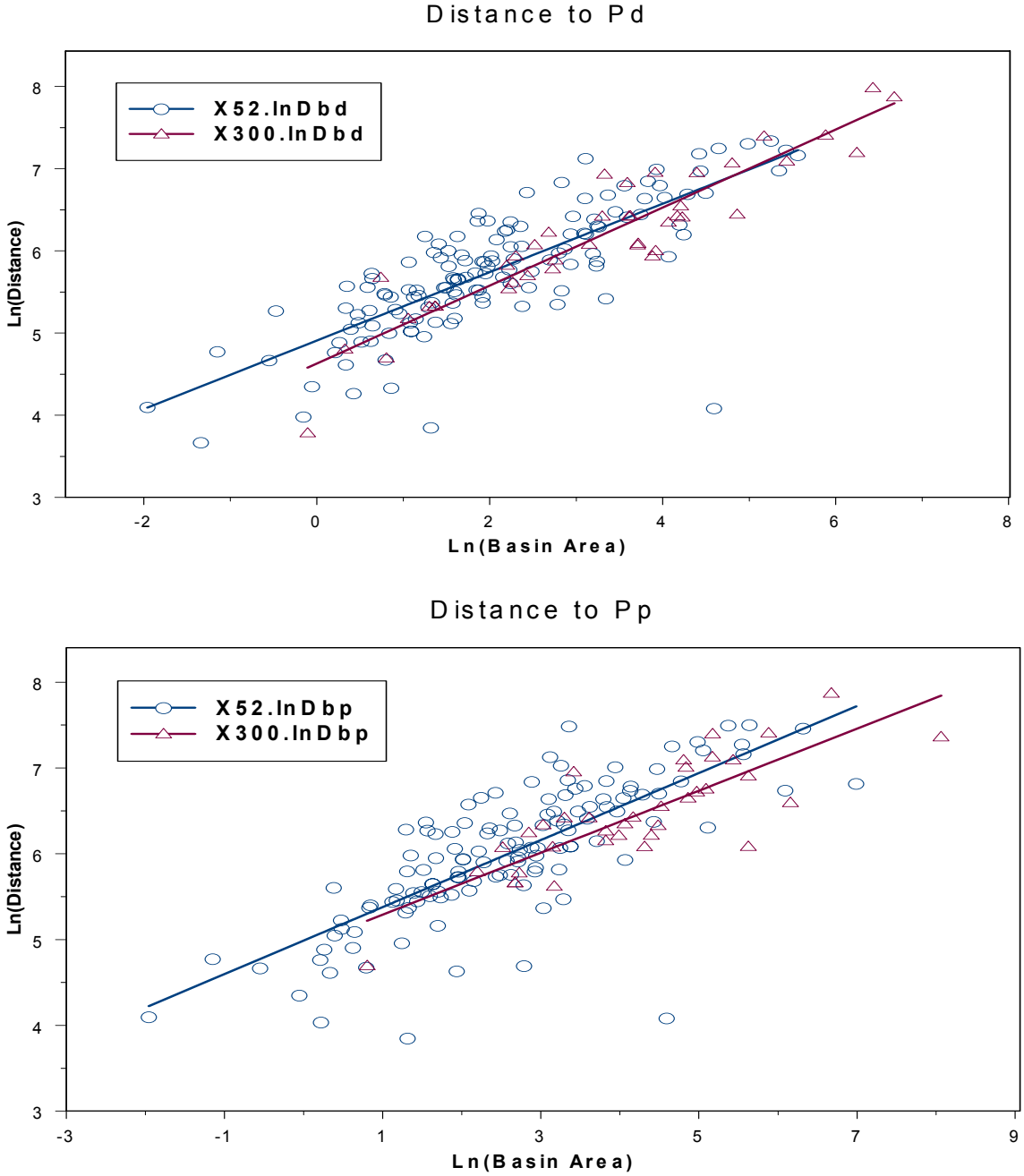


Figure 4. **Scatter plot of basin area and distance for distance to Pd and Pp showing basin areas separately. Both measurements are log-transformed.**

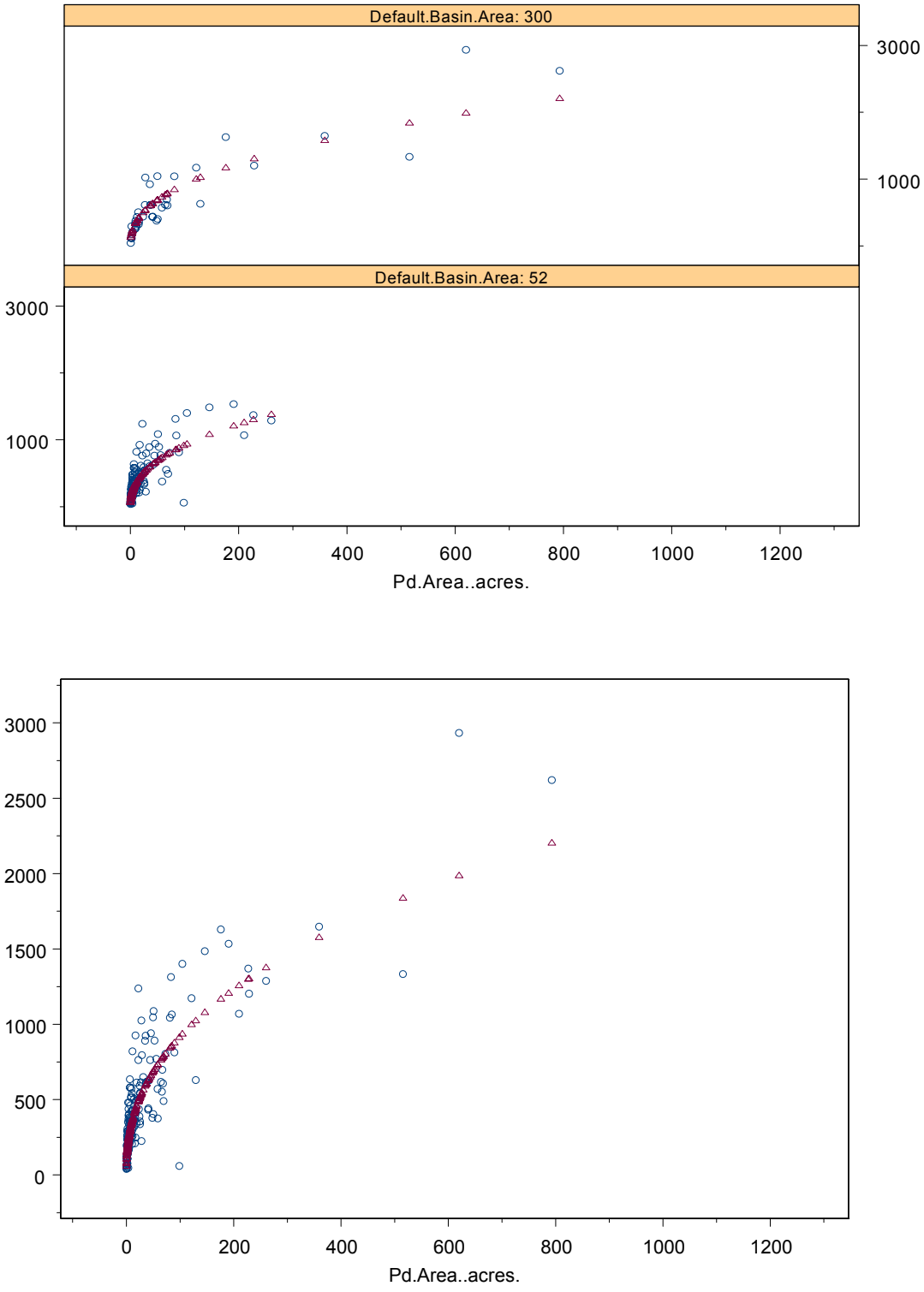


Figure 5. **Estimated relationship (estimated = Δ, observed data = o) between distance from divide to points P_d and basin areas (A_d) on the original scale by region and statewide.**

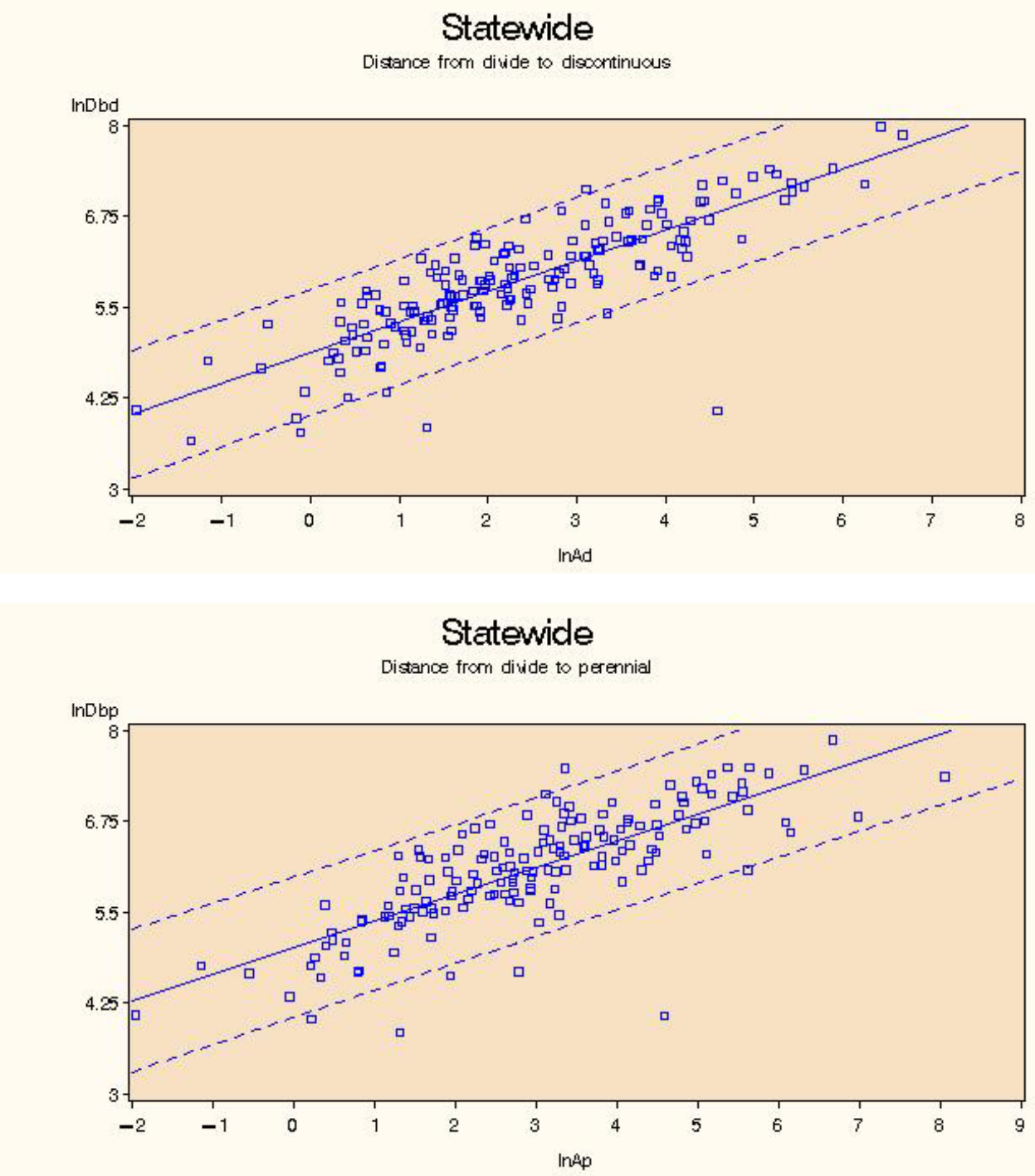


Figure 6. Relationship between distance from divide to points P_d and P_p and basin areas (A_d and A_p) for log-transformed data with prediction intervals.